

REMARKS/ARGUMENTS

Claims 1-30 are pending. No claims have been amended. No claims have been added. No claims have been cancelled. In view of the following arguments, withdrawal of all outstanding rejections and objections to pending claims 1-30 is respectfully requested.

Claim Rejections Under 35 USC §103(a)

Claims 1-4, 6-7, 9, 11-14, 16-17, 19, 21-24, 26-27, and 29 stand rejected under 35 USC §103(a) as being unpatentable over U.S. Patent No. 5,864,845 to Vorhees et al ("Vorhees"). This rejection is traversed.

The Voorhees Reference

Voorhees at col. 2, lines 23-65, teaches "facilitating World Wide Web searches by combining search result documents, as provided by separate search engines in response to a query, into one single integrated list so as to produce a single document with a ranked list of pages". In part this is accomplished by "training the computer for each search engine by clustering training queries and building cluster centroids. [T]he set of queries for which relevance data is known is called the training queries" (col. 2, lines 46-49). Voorhees at col. 3, lines 26-40, explicitly teaches that the training queries are clustered based on term similarity, and a select set of queries, those determined to have the highest measure of term similarity, are sent to one or more search engines. Voorhees, at col. 3, lines 41-50, teaches that relevant document distributions resulting from submission of the select set of queries to the search engines are determined from the search results. With respect to "relevant document distributions", Voorhees at col. 2, lines 40-49

1 at col. 2, lines 40-49, teaches that a relevant document is a document that was
2 previously indicated by a user as being relevant, wherein the document was
3 returned to the user responsive to submission of one of k training queries to a
4 search engine.

5 For each search engine's results and for that search engine's associated
6 calculated relevant document distributions, Voorhees at col. 4, lines 51-56, teaches
7 that a weight is assigned to the corresponding query cluster that was submitted to
8 the search engine – the weight being a measure of how effective “queries in the
9 cluster are for that search engine.” This is emphasized at col. 5, lines 51-53,
10 wherein Voorhees teaches “[a]ssign weights to each cluster reflecting the number
11 of relevant pages expected to be obtained by this search engine for queries similar
12 to those in the cluster.”

13 At col. 5, lines 57-67, Voorhees teaches “[t]o process an incoming query,
14 for each search engine[:] [f]ind the cluster centroid that is most similar to the
15 [incoming / current] query”, “[c]reate a query vector for the current query in the
16 vector space of the training queries”, “[c]ompute a vector similarity measure (e.g.,
17 the cosine) between the current query vector and each of the centroids”, “[c]hoose
18 the centroid that has the greatest similarity [‘[t]he assumption is that if two queries
19 retrieve many documents in common they are about the same topic’ (col. 4, lines
20 45-46)], and “[r]eturn the weight associated with the selected cluster as the weight
21 of the current search engine” and “[a]pportion the N slots in the retrieved set
22 according to the weights returned by each search engine.”

23 In view of the above, Voorhees teaches that training queries are clustered
24 based on term similarity, *not relevance data*. The fact that relevance data exists
25 for any particular query *merely qualifies the particular query to be clustered based*

1 on term similarity with other queries for which relevance data exists. A query
2 with corresponding relevance data is needed by Voorhees because, after training
3 queries clustered based on similar terms are submitted to a search engine, the
4 relevance data can be used in view of a new query submitted by an end-user to
5 evaluate the quality of each search engine's results as a function of topic area.

6 The Claimed Features

7 Claim 1 recites "identifying a same document and/or a plurality of similar
8 documents selected by a user in response to a plurality of queries", and
9 "responsive to identifying the same document and/or the similar documents,
10 generating a query cluster to indicate that the queries are similar independent of
11 whether individual ones of the queries comprise similar composition with respect
12 to other ones of the queries." Nowhere does Voorhees teach or suggest these
13 recited features.

14 In addressing claim 1, The action at page 3, admits that that Voorhees does
15 not teach or suggest "generating a query cluster to indicate that the queries are
16 similar independent of whether individual ones of the queries comprise similar
17 composition with respect to other ones of the queries", as claim 1 recites.
18 Additionally, the Action admits that Voorhees teaches clustering queries based on
19 similar query composition (page 2, section 2), which is contrary to what claim 1
20 recites. However, even in view of these concessions and without referring to
21 another reference in combination to supply these admittedly missing features, the
22 Action still concludes that the recited feature is obvious in view of Voorhees.
23 More particularly, the Action points to Voorhees teaching that "[s]earch engines
24 such as Excite and Alta Vista provide a query interface to information in these
25

1 pages [of Voorhees]". In view of this, the Action points out that queries input into
2 such a query interface (i.e.,) have no relationship to one another, and are thus,
3 independent from one another. In view of this reasoning, the Action concludes
4 that "it would have been obvious [...] to include queries are similar independent of
5 whether individual ones of the queries comprise similar composition in the system
6 of Voorhees" because "[g]enerating query clusters from random (each query
7 input is independent of each other) queries, improves search performance by
8 grouping similar queries." This conclusion is unsupportable.

9 The portion of Voorhees that the Action relies on for this modification to
10 Voorhees indicates, at most, that a search engine provides an interface for search
11 query input (col. 1, lines 47-59). Applicant respectfully submits, that just because
12 a user may submit separate queries to a search engine does not mean that
13 Voorhees now clusters queries independent of term similarity across the queries.
14 As already conceded by the Action, Voorhees teaches query clustering based on
15 similar query terms. This is very different from "identifying a same document
16 and/or a plurality of similar documents selected by a user in response to a plurality
17 of queries", and "responsive to identifying the same document and/or the similar
18 documents, generating a query cluster to indicate that the queries are similar
19 independent of whether individual ones of the queries comprise similar
20 composition with respect to other ones of the queries", as Applicant claims. This
21 claimed feature results in queries that are clustered as a function of a user's
22 judgment as to which documents are relevant to a query – not queries clustered
23 based on similar terms as Voorhees explicitly teaches.

24 In view of the above, a system of Voorhees may never "responsive to
25 identifying the same document and/or the similar documents, generating a query

1 cluster to indicate that the queries are similar independent of whether individual
2 ones of the queries comprise similar composition with respect to other ones of the
3 queries", as claim 1 recites.

4 For these reasons alone, the 35 USC §103(a) rejection of claim 1 over
5 Voorhees is improper and should be withdrawn.

6 Additionally, Applicant respectfully submits, that the proposed
7 modification of Voorhees to generate random query clusters independent of term
8 similarity, as suggested by the Action at page 3, changes the principal operation of
9 Voorhees. According to the MPEP §2143.01, "[i]f the proposed modification or
10 combination of the prior art would change the principle of operation of the prior
11 art invention being modified, then the teachings of the references are not sufficient
12 to render the claims prima facie obvious." In view of this the Office is urged to
13 review Voorhees col. 3, lines 26-31, and col. 5, line 57 through col. 6, line 20,
14 wherein Voorhees explicitly teaches that queries clustered by term similarity are
15 needed for proper operation of Voorhees search engine query clustering fusion
16 strategy. Thus, the proposed modification to Voorhees to generate clusters at
17 random and independent of term similarity would likely result in a similarity
18 measure between terms of a newly input user query and cluster centroids content
19 would be of no inherent value to Voorhees.

20 For this additional reason, the 35 USC §103(a) rejection of claim 1 in view
21 of Voorhees is improper and should be withdrawn.

22 Moreover, since the Action does not combine Voorhees with any secondary
23 reference to provide the missing features of claim 1, the Office is seemingly
24 relying on personal knowledge to arrive at the claimed features. "When a
25 rejection in an application is based on facts within the personal knowledge of an

1 employee of the office, the data shall be as specific as possible, and the reference
2 must be supported, when called for by the applicant, by the affidavit of such
3 employee, and such affidavit shall be subject to contradiction or explanation by
4 the affidavits of the applicant and other persons." 37 CFR §1.104(d)(2).

5 Accordingly, if this rejection of claim 1 is maintained on a similar basis in a
6 subsequent action, Applicant respectfully requests the Examiner to supply such an
7 affidavit to support these modifications to Voorhees.

8 Claims 2-10 depend from claim 1 and are allowable over Voorhees by
9 virtue of this dependency. For this reason alone, the 35 USC §103(a) rejection of
10 claims 2-10 are improper and should be withdrawn.

11 Additionally, claims 2-10 include other features that are not taught or
12 suggested by Voorhees. For example, claim 6 recites "constructing a thesaurus
13 comprising a plurality of synsets, wherein each synset comprises one or more
14 query clusters." In addressing the claimed "thesaurus" and "synsets", the Action
15 points to Voorhees col. 4, lines 38-60, col. 2, lines 23-49, col. 5, lines 13-67, and
16 col. 6, lines 1-19. These portions of Voorhees have been reviewed and do not
17 provide any teaching or suggestion of a "thesaurus" or a "synset", as Applicant
18 claims.

19 For this additional reason, the 35 USC §103(a) rejection of claim 6 should
20 be withdrawn.

21 It is respectfully requested that if claim 6 is again rejected in view of
22 Voorhees, that the Office particularly point out where Voorhees teaches or
23 suggests "a "thesaurus" or a "synset", as Applicant claims.

24 In another example, claim 9 recites "determining the similar documents
25 based on a hierarchical positioning between individual ones of a plurality of

1 documents commonly selected across the queries." In addressing this feature, the
2 Action points to Voorhees at col. 4, lines 38-60, col. 2, lines 23-49, col. 5, lines
3 13-67, and col. 6, lines 1-19. With respect to use of a hierarchy, Voorhees at most
4 teaches "form clusters from hierarchy by considering all queries that cluster above
5 a certain threshold to belong to the same cluster." Applicant respectfully submits,
6 that forming a query cluster as taught by Voorhees results in a query cluster –
7 nothing more. In contrast, "determining the similar documents based on a
8 hierarchical positioning between individual ones of a plurality of documents
9 commonly selected across the queries" results in a similar document
10 determination. Clearly, this teaching of forming a query cluster by Voorhees does
11 not teach or suggest the features of claim 9.

12 For this additional reason, the 35 USC §103(a) rejection of claim 9 should
13 be withdrawn.

14 Claim 11 recites "identifying a same document and/or a plurality of similar
15 documents selected by a user in response to a plurality of queries", and
16 "responsive to identifying the same document and/or the similar documents,
17 generating a query cluster to indicate that the queries are similar independent of
18 whether individual ones of the queries comprise similar composition with respect
19 to other ones of the queries." For the reasons already discussed, Voorhees does
20 not teach or suggest these recited features.

21 Accordingly, the 35 USC §103(a) rejection of claim 11 is improper and
22 should be withdrawn.

23 Claims 12-20 depend from claim 11 and are allowable over Voorhees by
24 virtue of this dependency. For this reason alone, the 35 USC §103(a) rejection of
25 claims 12-20 are improper and should be withdrawn. Additionally, claims 12-20

1 include other features that are not taught or suggested by Voorhees for the reasons
2 already discussed.

3 **Claim 21** recites "identifying a same document and/or a plurality of similar
4 documents selected by a user in response to a plurality of queries", and
5 "responsive to identifying the same document and/or the similar documents,
6 generating a query cluster to indicate that the queries are similar independent of
7 whether individual ones of the queries comprise similar composition with respect
8 to other ones of the queries." For the reasons already discussed, Voorhees does
9 not teach or suggest these recited features.

10 Accordingly, the 35 USC §103(a) rejection of claim 21 is improper and
11 should be withdrawn.

12 **Claims 21-30** depend from claim 20 and are allowable over Voorhees by
13 virtue of this dependency. For this reason alone, the 35 USC §103(a) rejection of
14 claims 21-30 are improper and should be withdrawn. Additionally, claims 21-30
15 include other features that are not taught or suggested by Voorhees for the reasons
16 already discussed.

17
18 **Claim Objections**

19 Claims 5, 8, 10, 15, 18, 20, 25, 28, and 30 stand objected to as being
20 dependent on a rejected base claim, but would be allowable if re-written in
21 independent form including all limitations of the base claim and any intervening
22 claims. Applicant thanks the Office for this indication of allowability. However,
23 it is respectfully submitted that claims 5, 8, 10, 15, 18, 20, 25, 28, and 30, as well
24
25

1 as their respective base claims and any intervening claims, are patentably
2 distinguished over the references of record for the reasons discussed above.

3
4 **Conclusion**

5 Claims 1-30 are in condition for allowance and action to that end is
6 respectfully requested. Should any issue remain that prevents allowance of the
7 application, the Office is encouraged to contact the undersigned prior or issuance
8 of a subsequent Office action.

9
10
11 Respectfully Submitted,

12
13 Dated: 4/29/2004

14 By: Brian Hart
15 Brian G. Hart
16 Reg. No. 44, 421
17 (303) 539-0265 x 241
18
19
20
21
22
23
24
25